

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 13 日現在

機関番号：32665

研究種目：若手研究(B)

研究期間：2014～2015

課題番号：26730183

研究課題名(和文)金管楽器演奏に対する音響空間と筋電空間の相互マッピング

研究課題名(英文) Mutual Mapping between Acoustic Space and Muscle Activity Space in Playing Brass Instrument

研究代表者

北原 鉄朗 (KITAHARA, Tetsuro)

日本大学・文理学部・准教授

研究者番号：00454710

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：本研究では、金管楽器演奏時の音から筋活動(もしくはその逆)を予測する技術の研究開発を通じて、楽器練習に科学的妥当性をもたらすことである。主に次の成果が得られた。

- (1) 音の高さや強さを様々に変えてトランペットを演奏した際の筋活動を分析した結果、音が高くなるにつれて筋活動が活発になることが分かった。
- (2) 演奏時の音響信号から抽出した特徴量から筋活動量への変換を試みた結果、線形性を仮定した手法(アフィン変換)よりも線形性を仮定しない手法(ニューラルネットワーク)の方が精度が高いことが分かった。

研究成果の概要(英文)：The purpose of this research is to develop a technique for predict the muscle activities from acoustic features (and vice versa) in playing the trumpet. We obtained the following results:

- (1) The muscle activity is greater as the pitch is higher.
- (2) In transforming the acoustic features extracted from the sound of playing the trumpet to muscle activity values, a non-linear method (neural network) is more accurate than a linear method (affine map).

研究分野：音楽情報処理

キーワード：筋活動 金管楽器 変換

1. 研究開始当初の背景

楽器演奏は、生涯にわたって楽しむことができる娯楽 1 つである。練習の反復により自らのスキルアップを実感し、他者と一緒にアンサンブル演奏を楽しむことで社会的なつながりを促進する。そのため、楽器演奏スキルを修得することに憧れを抱く者は多い。しかし、基本的なスキルの修得には長期の反復練習が不可欠であり、挫折する者も少なくない。特に、金管楽器は、バイオリンなど並んで音を出すこと自体が難しい楽器として知られている。金管楽器の演奏には、

- (1) 口の周囲の筋肉を適度に緊張させ、独特の口の形を作ってマウスピースに押し付ける、
- (2) 唇を振動させる、
- (3) 呼気を吹き込む、

といった 3 つの動作を、音を聴いて微調整しながら同時に行う必要がある。この独特の口の形は「アンブシュア」と呼ばれ、この形状を一定に保つことが重要であると言われている（伊藤他、電気学会論文誌、2011）。しかし、周囲の筋肉に余計な力が入ってくると、疲れやすく長時間の演奏に支障を来たすだけでなく、アンブシュアを一定に保つことが難しくなり、「ピッチが安定したロングトーンを出せない」「きれいに高音が出ない」といった問題が生じる。しかし、学習者が自ら余計な力みに気づくのは難しく、科学的知見と情報技術を用いて適切な支援を行う技術の確立が望まれる。

2. 研究の目的

本研究の目的は、金管楽器演奏における音響的特徴と筋活動の測定値とを相互に変換する技術の開発を通じて、楽器演奏の基礎練習に「科学的妥当性」「手軽さ」「楽しさ」を与えることにある。楽器演奏は、生涯楽しめる娯楽であるが、金管楽器のように音を出すこと自体が難しい楽器では、素人による独自の練習により無理のある演奏法を身につけてしまい、その後の上達に支障を来たす可能性がある。本申請の研究期間では、上述の目的を実現するための基盤技術と

して、金管楽器演奏中の筋活動を音から予測（およびその逆）をすることにより、練習者が自らの筋活動をリアルタイムにモニタリングできる技術の確立を目指す。

3. 研究の方法

金管楽器のうちトランペットに着目し、トランペット演奏時の音響信号と筋活動を同時に計測する。次に、音響信号から基本周波数やパワーなどの特徴量を抽出し、特徴空間を構成する（音響空間とよぶ）。筋活動については 2 チャンネルの筋電センサを口唇周囲に取り付け、そこから得られる 2 次元の筋電値を平滑化したものを用いる（筋電空間とよぶ）。音響空間から筋電空間への写像（またその逆）を構成することで、音からの筋活動の予測（およびその逆）を実現する。

音響空間から筋電空間への写像は、非線形かつ時間的な依存関係が存在すると思われる。このような、非線形かつ時間的な依存関係が存在する写像を実現するには recurrent neural network (RNN) を用いるのが望ましいと思われる。しかし、こうした試みが過去になかったことに鑑み、できるだけシンプルな方法、つまり、線形な変換、時間的な依存関係を考慮しない変換だけでどの程度の精度が得られるかを確かめるべきである。そこで、まずは、線形な変換手法として「アフィン変換」を用い、その後、非線形な変換手法（時間的な依存関係は考慮しない）としてニューラルネットワークを用いる。

4. 研究成果

(1) 筋活動の調査実験

音響空間と筋電空間の相互変換の試行に先立ち、音の高さや強さを変えてトランペットを演奏した際にどのように筋活動が変化するかを調査した。White ら (1974) は、同様の調査をしており、次の結論を出している。

- 低音域の演奏中よりも高音域を演奏

中の方が、筋活動が高まる。低強度の演奏中よりも高強度の演奏中の方が、筋活動が高まる。ただし、強度よりも音域の差の方が筋活動へ与える影響は大きい。

- 初心者は上唇の筋活動が下唇の筋活動より小さく、熟達者にその違いはない。

一方、後藤田らは、自身の研究で口輪筋などの咀嚼筋は音量の大小で各筋の活動量に明らかな変化はなかったと結論付けている。

そこで、Whiteと同様の条件で実験を行った。計測機器は、筋電計測に「ワイヤレス EMG ロガー(ロジカルプロダクト製)」と「筋電センサー(追坂電子機器製)」を使用し、録音に「C519M (AKG 製)」を使用した。また、楽器による計測誤差を無くするため、全員「YTR-4335GS (YAMAHA 製)」を使用した。筋活動の測定場所は、Whiteらと同じく

- 上唇の口輪筋、orbicularis oris superioris (OOS)
- 下唇の口輪筋、orbicularis oris inferioris (OOI)
- 口角下制筋、depressor anguli oris (DAO)
- 口角挙筋、levator anguli oris (LAO)

とした。筋電図と楽器音を同時に記録し、楽器音が鳴っている際の最大振幅の 75%以上の区間を発音区間と定義し、その部分に相当する筋電図を利用する。利用するには RMS 値を計算し、使用した。被験者は Table 1 の 4 名である。

Table 1 被験者

被験者	年齢	性別	楽器歴
被験者 A	22 歳	男性	トランペット 歴 7 年
被験者 B	22 歳	女性	ホルン 歴 12 年 (トランペット 歴無し)
被験者 C	23 歳	男性	ホルン 歴 8 年 (トランペット 歴無し)
被験者 D	19 歳	女性	トランペット 歴 半年

【実験1】

低音域の演奏中よりも高音域を演奏中の方が、筋活動が高まるという仮説について検証し

ていく。被験者には B♭3 から変口長調の音階で 1 音ずつ口を休み、演奏可能な音高まで吹いてもらった。低音域には、B♭3 から 3 音、高音域には、はっきりと発音された最も高い音を含む 3 音を使用して比較検定した。検定にはマンホイットニーの U 検定を用い、2 つの群が等しい中央値を持つかどうかを有意水準 5% で検定している。被験者個人で最高音が違うため、

- 被験者個人の音域内で低音域と高音域の筋活動を検定した場合
 - 被験者全員の高音域部分を揃えて低音域と高音域の筋活動を検定した場合
- の 2 パターンについて検証した。

被験者個人の音域に着目し、それぞれの音域内で低音域、高音域を比較した場合、Table 2 の結果となり、全被験者、各筋肉において有意水準 5% で棄却できる結果となった。これにより、低音域の演奏中よりも高音域を演奏中の方が、筋活動が高まると言える。また、被験者全員の音域を B♭3 から B♭4 までの 1 オクターブ分に揃えて比較した場合、Table 3 の結果となり、棄却できない箇所が表れた。これにより、特定の音高に対して演奏に必要な筋活動量が決まるのではなく、各被験者の演奏可能な音域によって、低音域と高音域での筋活動に差が出るという事が分かる。

Table 2 被験者個人の低音域、高音域を比較した場合

P 値	被験者 A	被験者 B	被験者 C	被験者 D
OOS	5.3111×10^{-7}	2.2422×10^{-6}	3.8144×10^{-7}	7.3666×10^{-7}
OOI	5.6062×10^{-6}	3.8144×10^{-7}	3.8144×10^{-7}	4.7810×10^{-5}
DAO	1.2000×10^{-3}	4.1135×10^{-5}	4.1135×10^{-5}	8.2271×10^{-5}
LAO	2.8000×10^{-3}	4.1135×10^{-5}	4.1135×10^{-5}	8.2271×10^{-5}

Table 3 被験者全員の音域を揃え比較した場合

P 値	被験者 A	被験者 B	被験者 C	被験者 D
OOS	5.0050×10^{-4}	2.2422×10^{-6}	1.9206×10^{-6}	5.4551×10^{-5}
OOI	2.7800×10^{-2}	3.8144×10^{-7}	3.2278×10^{-7}	2.5000×10^{-3}
DAO	3.4010×10^{-1}	4.1135×10^{-5}	4.1135×10^{-5}	7.700×10^{-2}
LAO	7.9620×10^{-1}	4.1135×10^{-5}	4.1135×10^{-5}	4.9362×10^{-4}

【実験2】

低強度の演奏中よりも高強度の演奏中の

方が、筋活動が高まるという仮説について検証する。被験者には低強度と高強度をそれぞれ B \triangleright 3、B \triangleright 4 の音高で演奏してもらい、被験者ごとの、低強度演奏時と高強度演奏時の筋活動を比較検定した。検定には実験 1 と同様、U 検定を用いた。

被験者個人で比較し、Table 4 の結果が得られた。これにより、OOS, OOI の両筋肉に関しては、全被験者共通で有意水準 5% で有意差ありとの判定となった。しかし、LAO, DAO に関しては、有意差ありの被験者 (A,B) も居れば、有意差無しの被験者 (C,D) も居た。また、音域と強度の筋活動への影響度を調べるため、音域間の筋活動値の差と強度間の筋活動値の差を求め、U 検定で比較した場合、Table 5 のような結果となり、被験者ごとの各筋肉によって有意差ありの場合と有意差無しの場合があった。よって、強度よりも音域の方が、筋活動に影響を与えると言えない。

Table 4 低強度演奏時と高強度演奏時の筋活動を比較

P 値	被験者 A	被験者 B	被験者 C	被験者 D
OOS	4.7768×10^{-4}	1.7900×10^{-2}	1.7900×10^{-2}	2.9000×10^{-3}
OOI	4.6949×10^{-5}	3.8420×10^{-4}	4.0400×10^{-2}	9.0094×10^{-4}
LAO	1.5200×10^{-2}	2.6000×10^{-2}	1.7970×10^{-1}	8.4400×10^{-2}
DAO	2.2000×10^{-3}	2.2000×10^{-3}	1.7970×10^{-1}	6.4900×10^{-2}

Table 5 音域と強度の筋活動への影響について

P 値	被験者 A	被験者 B	被験者 C	被験者 D
OOS	1.2600×10^{-1}	7.3400×10^{-2}	7.3000×10^{-3}	5.4440×10^{-1}
OOI	4.2733×10^{-4}	2.6020×10^{-1}	3.6442×10^{-5}	2.1450×10^{-1}
LAO	5.8870×10^{-1}	9.3720×10^{-1}	2.2000×10^{-3}	4.1100×10^{-2}
DAO	2.6000×10^{-2}	1.5200×10^{-2}	2.2000×10^{-3}	5.8870×10^{-1}

【実験3】

初心者は上唇の筋活動が下唇の筋活動より小さく、熟達者にその違いはないという仮説について検証する。各被験者の B \triangleright 3 から B \triangleright 4 までの音を使用し、OOS と OOI の筋活動を比較検定した。検定には実験 1 と同様、U 検定を使用した。

各被験者の OOS および OOI について検定した結果、Table 6 の結果が得られた。これにより、トランペット歴の長い被験者 A は OOS と OOI の筋肉間に有意差は無く、ト

ランペット歴の浅い他の被験者は有意差がある事が分かる。

Table 6 OOS および OOI について検定した結果

被験者 A	被験者 B	被験者 C	被験者 D
2.4400×10^{-1}	2.5000×10^{-3}	6.2096×10^{-8}	3.0487×10^{-9}

(2) 音響空間から筋電空間への変換実験

トランペット演奏時の音響信号と筋活動量を同時に測定し、その関係性を学習することで、音響信号から筋活動量を予測することができるようになる。簡単のため、1 つ 1 つの音符を切って発音するものとし、個々の音符に対応する音響信号を切り出して、その安定区間の音響信号および筋活動量のみを用いる。

演奏音のデジタル音響信号 (サンプリング周波数: 48kHz) が与えられると、振幅が最大振幅の 10% 以下の箇所を無音とみなし、無音の箇所で音響信号を分割する。分割された各音響信号 (個々の音符に対応するとみなす) に対して振幅が最大振幅の 75% 以上の区間を安定区間とみなす。安定区間から 10ms 毎に基本周波数 (F0)、振幅 (Amp)、スペクトル重心 (SC)、スペクトルロールオフ (SR)、スペクトルフラックス (SF) MFCC (13 次まで) を抽出する。F0 の抽出には yeg-nanarayana の手法、SC, SR, SF, MFCC の抽出には MIRtoolbox を用いる。

演奏音と同時に測定した上下口輪筋の筋活動のデジタル信号 (サンプリング周波数: 1kHz) が与えられると、演奏音の安定区間に対応した筋活動値を抽出する。その筋活動値から 10ms 毎に RMS 値を算出する。

次に、音響特徴量 x_1, \dots, x_n を筋活動量 y_1, y_2 へ変換する。変換手法として、両者の関係に線形性を仮定したアフィン変換 (AT)、線形性を仮定しないニューラルネットワーク (NN) を用いる。ニューラルネットワークの中間ノードは 2 個とする。

トランペット演奏歴 7 年の熟達者に、変口

長調の音階を 1 音ずつ口を休めながら B♭3 から演奏可能な最高音 (C6)まで演奏してもらった。演奏音の収録には「C519M (AKG 製)」、筋活動計測には「ワイヤレス EMG ロガー (ロジカルプロダクト製)」および「筋電センサー (追坂電子機器製)」を用いた。同じ条件で 2 度演奏してもらい、1 回めのデータを学習用に、2 回めのデータを評価用に使用した。

AT および NN を用いて筋活動量を音響特徴量から推定した結果、および筋活動量の実測値を Fig. 1, Fig. 2, Table.1 に示す。紙面の都合により上唇の筋活動のみ示す。Fig. 1 では、closedで低音域から中音域までの変換が上手くいっているが高音域での変換が上手くいっていないことがわかる。そのため open で変換した際に、高音域での誤差が大きくなってしまっている。また、Fig. 2 では closedでの変換は低音域から高音域まで上手くいっているのがわかる。Table 1 の AT に着目すると全ての音響特徴量を上下の筋活動量へ変換した場合よりも SR 以外で変換した場合の誤差が最も小さくなり、SC 以外で変換した場合の誤差が最も大きくなった。また AT に比べて NN の方が、全ての場合において誤差を約半分に抑えられていることがわかる。

AT と NN はともに、上唇よりも下唇の方が誤差が高くなっていて、下唇よりも上唇の方が今回の変換に適していると考えられる。AT において、SR 以外で変換した場合の誤差が最も小さく、SC を抜いて変換した場合の誤差が最も大きくなったことについて、SR は今回の変換で有効な特徴量ではなく逆に SC は有効な特徴量であることが考えられる。

評価に用いた音響特徴量を図示すると Fig. 3 のようになり、SR と SC が中音域終盤から数値が高くなる部分がある。学習に用いた音響特徴量にはその特徴が見られず、結果として誤差が高くなってしまったと考えられる。この SR と SC が高くなる特徴は、他の被験者にも

見られる現象なのか、なぜ高くなるのかを今後詳しく調査する必要がある。また、SR と SC の間に相関関係がみられることから、今回の変換では SC が SR より影響して誤差が大きくなったと考えている。

NN において、上唇では F0 以外で変換した場合の誤差が最も小さく MFCC以外で変換した場合の誤差が最も大きくなった。また、下唇では MFCC以外で変換した場合の誤差がもっとも小さく、SF 以外で変換した場合の誤差が最も大きくなり、上唇と下唇とで変換に有効な音響特徴量が違うことがわかる。

AT と NN で変換に影響を与えた音響特徴量が違うことについて、学習時に影響した音響特徴量が違うためだと考えられる。今後はどの音響特徴量が変換に最も影響を与えるかを、各音響特徴量だけでの変換を行い調査する必要がある。NN による推定 (Fig. 2)では瞬間的に誤差が大きくなっている箇所があった。これは、F0 推定において倍ピッチエラーや半ピッチエラーが生じている (Fig. 3 左上) からであると考えられる。

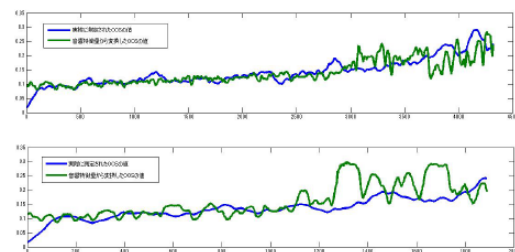


Fig. 1 AT による上唇の推定値と実測値 (上: closed, 下: open)

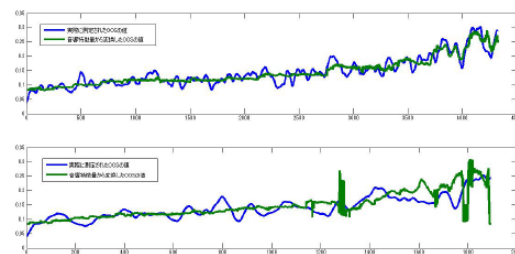


Fig. 2 NN による上唇の推定値と実測値 (上: closed, 下: open)

Table 1 音響特徴量から筋活動量への変換の二乗平均誤差 ($\times 10^3$)

用いた特徴量	上唇		下唇	
	AT	NN	AT	NN
すべて	2.59	1.43	7.23	4.07
F0 以外	2.73	0.89	7.54	3.26
Amp 以外	2.52	1.78	7.16	3.57
SC 以外	2.97	1.10	8.00	3.04
SR 以外	1.68	1.16	5.54	3.63
SF 以外	2.45	1.22	7.16	3.68
MFCC 以外	2.68	1.85	7.37	2.97

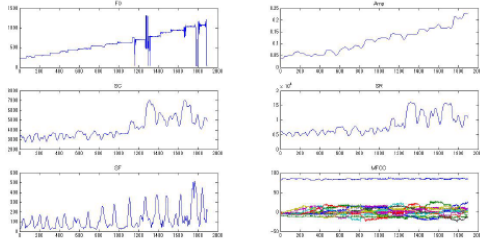


Fig. 3 評価時の変換に用いた音響特徴量

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 0件)

[学会発表](計 2件)

(1) 佐藤 愛, 北原 鉄朗, 寺澤 洋子, 松原 正樹 : “トランペット演奏時の口唇周囲および腹部の筋活動と音響的特徴の関係”, 日本音響学会 2016年春季研究発表会 講演論文集, 2-10-7, 桐蔭横浜大学(神奈川県), 2016年3月10日.

(2) Tetsuro Kitahara, Shogo Matsukata, Masaki Matsubara, and Hiroko Terasawa: "A Preliminary Experiment of Predicting Muscle Activity from Musical Acoustic Features", Proceedings of the 7th International Workshop on Machine Learning and Music (MML 2014), Barcelona (Spain), 28th November 2014.

[図書](計 0件)

[産業財産権]

○出願状況(計 0件)

○取得状況(計 0件)

[その他]

ホームページ等

<http://www.kthrlab.jp/>

6. 研究組織

(1) 研究代表者

北原 鉄朗 (KITAHARA, Tetsuro)

日本大学・文理学部・准教授

研究者番号 : 00454710

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

寺澤 洋子 (TERASAWA, Hiroko)

筑波大学・大学院図書館情報メディア

研究科・助教

松原 正樹 (MATSUBARA, Masaki)

筑波大学・大学院図書館情報メディア

研究科・特任助教